International Journal of
Medical Toxicology & Forensic Medicine
Winter 2020, Volume 10, Number 1

# Research Paper: The Homology Modeling and Docking Investigation of Human Cathepsin B

**Afshin Khara**[1], **Ehsan Jahangirian**[2*], **Hossein Tarrahimofrad**[2]

*1. Amin Police University, Tehran, Iran.*
*2. Department of Animal Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran.*

Use your device to scan and read the article online

**Citation:** Khara A, Jahangirian E, Tarrahimofrad H. The Homology Modeling and Docking Investigation of Human Cathepsin B. International Journal of Medical Toxicology and Forensic Medicine. 2020; 10(1):26687. https://doi.org/10.32598/ijmtfm.v10i1.26687

https://doi.org/10.32598/ijmtfm.v10i1.26687



**Article info:**
**Received:** 10 Aug 2019
**First Revision:** 15 Aug 2019
**Accepted:** 03 Sep 2019
**Published:** 03 Apr 2020

**Keywords:**
Cathepsins, Cysteine, Cysteine Proteinase Inhibitors

## ABSTRACT

**Background:** Cathepsin B comprises a group of lysosomal cysteine proteases belonging to the Papain family; it has an intracellular function in the process of protein catabolism, antigen processing in the immune response, and Alzheimer's disease. In cancers, cathepsin B interferes with autophagy and intracellular catabolism, and breaks down extracellular matrix, decreases protease inhibitors expression, and ultimately helps to accelerate metastasis, tumor malignancy, and reduce immune resistance.

**Methods:** In this study, the 3D structure of cathepsin B was constructed using modeler and Iterative Threading ASSEmbly Refinement (I-TASSER), based on similarity to the crystallographic model of procathepsin B (1PBH). Then, the predicted cathepsin B model was evaluated using PROCHECK and PROSA for quality and reliability. Molecular studies suggested that the amino acids cysteine 108, histidine 189, and histidine 190 form the envelope of the active site of cathepsin B. The docking studies of cathepsin B was performed with protease inhibitors cystatin C, E-64 and leupeptin.

**Results:** The lowest binding energy was related to the cathepsin B-E-64 complex. Accordingly, it was found that E64 interacts with the amino acid cysteine 108 of the active site of cathepsin B. Leupeptin made 2 hydrogen bonds with cathepsin B, but none with the active site of cathepsin amino acids. Cystatin C established a hydrogen bond with the arginine 18 of cathepsin B and made electrostatic bonds with aspartate 148 of cathepsin B.

**Conclusion:** Therefore, the bioinformatics and docking studies of cathepsin B with its inhibitors could be used as reliable identification, treatment, and alternative methods for selecting the inhibitors and controllers of cancer progression.

*** Corresponding Author:**
***Ehsan Jahangirian, PhD.***
***Address:*** *Department of Animal Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran.*
***Tel:*** *+98 (912) 2076468*
***E-mail:*** *ehsan.jahangirian@yahoo.com*

## 1. Introduction

Cathepsin B is a member of Papain family; with its hydrolase activity as a proprietary enzyme with a molecular weight of 38 kDa, it is classified as a cysteine protease. Cathepsin B has V-cleft catalytic active site surrounded in a cleft by tow domains. Most cathepsins have a cysteine residue at their active site that interacts with histidine or aspartate to form a tetrad catalytic site. The nucleophilic attack is conducted by the sulfur amino acid agent; cysteine acts as a proton donor at this step and eventually leads to the cleavage of the peptide bonds. Cathepsin B has endopeptidase and exopeptidase activities [1]. Cathepsin B plays an essential role within the cell, as it helps to renovation proteins, the inflammatory response to antigen, and tissue regeneration. However, when excreted, it plays a destructive role, like the Extracellular Matrix (ECM) degradation [2].

The attenuation of the extracellular matrix enhances tumor regeneration, angiogenesis, tumor migration, invasion, and even cancer metastasis. Accordingly, by determining the serum levels of cathepsins L and B, we could detect the extent of cancer progression [3]. Cathepsin B, currently used as a prognostic marker (tumor marker) for cancer, could be an appropriate target for developing anticancer medications [4]. Understanding the functional mechanisms and binding sites and investigating how to inhibit cathepsin B activity using bioinformatics software is a very viable and justifiable solution in this respect. Numerous inhibitors for cysteine proteases have been identified.

Cystatin C is a class of natural cysteine protease inhibitors that decreases its ratio to cathepsin B during cancer; however, as increases cystatin C expression, it increases cathepsin B activity inhibition [5, 6]. Increased cystatin C expression alters the metastatic properties of melanocyte cells. It also reduces their mobility and aggression. For example, the serum levels of cystatin C/cathepsin B complex in patients with lung cancer are lower than those of healthy individuals [7]. This is because cathepsins also have hydrolase activity similar to cysteine proteases. Besides, they are inhibited by the synthetic inhibitors of cysteine proteases. A group of synthetic cysteine protease inhibitors, like E-64 and leupeptin as epoxide inhibitors, specifically inhibit cysteine protease activity. E-64 and leupeptin could inhibit cathepsin B activities [8, 9]. In this study, using docking studies, we investigated and compared the interactions between chemical and natural cysteine protease inhibitors with cathepsin B.

## 2. Materials and methods

Initially, online software http://web.expasy.org/protparam was used to calculate biochemical properties, including molecular weight, isoelectric point, instability index, aliphatic coefficient, and the net charge of cathepsin B [10]. The peptide signal was calculated by the online software http://www.cbs.dtu.dk/services/SignalP [11]. The possibility of disulfide bonding was investigated using http://disulfind.dsi.unifi.it website [12].

Next, Http://www.sbg.bio.ic.ac.uk/pHyre2 was used to predict the secondary structure of online software [13]. The server was defined by the Self-Optimized Prediction Method (SOPM). As a result, the increased rate of the second-order prediction improves based on the amino acid sequence.

Predicting the active site of the modeled protein was performed by combining data from two online software. Using the online website http://www.sbg.bio.ic.ac.uk/~3dligandsite/, the amino acids active site of cathepsin B were predicted [14]. Cathepsin B binding sites were also identified using the software https://zhanglab.ccmb.med.umich.edu/COFACTOR/ [15, 16]. Finally, by comparing the results of both online software, the amino acids involved in the active site were selected.

Using the Uniprot database, the blast results were stored at the 95% default similarity between the cathepsin B amino acid sequence and the amino acid sequence of other similar proteins. Then, using the Multiple Alignment Tool online website www.ebi.ac.uk, based on CLUSTAL O (1.2.4) algorithm, the conserved sequence of cathepsin B was evaluated and compared.

Modeller V.9 software was used to create the 3D building for homology modeling. The crystallographic structure of cathepsin B is only available as procathepsin; therefore, the cathepsin amino acid sequence was incorporated into the Blast-PDB. The 1G96 crystallographic structure that demonstrated the highest similarity and e-Value to cathepsin B was used for the Three-Dimensional (3D) structure of cathepsin B.

The websites http://services.mbi.ucla.edu/PROCHECK and http://services.mbi.ucla.edu/SAVES were used to verify the 3D predicted structure [17]. The collected data from WHAT_CHECK, ERRAT, VERIFY_3D, and CRYST1 record matches were also reviewed. The website https://prosa.services.came.sbg.ac.at/prosa.pHp was used to determine the Z-Score point and protein-energy balance [18, 19]. The Ramachandran plot graph was plotted by the online website

http://mordred.bioc.cam.ac.uk/~rapper/rampage.pHp [20]. The optimal spatial positioning of the protein was identified in terms of optimum energy by Chimera [21].

The mol2 structure of E-64 and leupeptin inhibitors was first obtained by PubChem. Then, the mol2 structure of E-64 and leupeptin together with the predicted 3-D structure of cathepsin B was corrected in Chimera for energy. The PDB structure of cathepsin B was implemented in Molegro Virtual Docker (MVD) software. All water molecules were removed from the PDB structure of cathepsin B. Nest, polar hydrogen atoms were added to the structure, and the amount and volume of the components were calculated by MVD. After preparing the cathepsin 3-D structure in the MVD, the cavity confine was determined.

The closest cavity to the active site of cathepsin B was selected and the grid size of the docking box was 12×16×10 with a central distance of Å 0.32. Docking was then performed between cathepsin B and E-64, as well as cathepsin B and leupeptin. At each step, the best complexes were selected in terms of binding energy. For protein-protein docking between cathepsin B and cystatin C, all step of cathepsin B preparation for cystatin C were also performed. Protein-protein docking between cathepsin B and cysteine C was performed using the ZDOCK server. H-bond formation was investigated by LIGPLOT [22].

## 3. Results

The molecular weight of cathepsin was estimated to be 37.82 kDa. Cathepsin B has an isoelectric point (pI: 5.88). The interpretation of cathepsin stability was determined using the instability index scale of 31.25. The Grand Average of Hydropathicity (GRAVY) was equal to -0.394. The aliphatic coefficient of the enzyme was measured as 64.69. The half-life of this protein in mammalian cells was estimated to be approximately 30 hours. Cathepsin B had 6 disulfide bonds due to its amino acid content.

The secondary structure of the protein was predicted using the SOPMA method. The relevant data revealed that the second structure of cathepsin B consisted of 24.8% alpha-helix, 16.8% beta-sheet, and 58.4% of the total random coil/beta-turn (Figure 1).

The multiple alignments of the cathepsin B amino acid sequence were performed using blast at the UniProt database. The results of multiple alignments suggested the conserved regions and cathepsin B active site amino acids with similar other proteins. As per Figure 2, the cysteine 108, histidine 189, and histidine 190 located in conserved regions were involved in the formation of the active site of cathepsin B.

The cathepsin B sequence was submitted in BLAST-PDB and the 1PBH structure was selected as the base structure of the homology modeling. As shown in Figure 3a, the Z-SCORE point was set at -6.34. In the graph, to determine the negative energy level of the protein, the amino acid sequence was considered lower than zero (Figure 3b). In Figure 3c, the red and blue spots of the protein indicate high-energy and low-energy areas, respectively. Ramachandran graph analysis relative to the quality of the predicted model indicated that 82.2% of amino acids were found in most favored regions, 15.6% in additional allowed regions, 1.1% in the areas of generously permitted regions, and only 1.1% were reported in disallowed regions (Figure 4). VERIFY-3D results for cathepsin B were found to be above 2 (Figure 5).

The PDB multiple alignments between the predicted cathepsin B model and the 5 predicted by the I-Tasser online software at http://ekhidna2.biocenter.helsinki.fi suggested an acceptable similarity between them (Figure 6). As per Figure 7, superimposition using Chimera (RMSD Å 2.32) revealed the active site of the cathepsin B structure consisting of the amino acids Cys 108, His 189, and His 190. By Chimera software, the best spatial positioning for enzyme optimal energy in the docks was obtained (Figure 8).

The cathepsin B-E-64 complex indicated that the nitrogen atom 4 of the E-64 ligand interacts with oxygen atom 8 of the cysteine 108 at the active site of cathepsin B at 2.68 angstroms by the hydrogen bond (Figure 9a). Additionally, the glutamine 115 of cathepsin B forms two electrostatic bonds with the nitrogen atom 7 of the E-64 ligand. Some of the cathepsin B amino acids also interact with E-64 by the citric bond (Figure 9b). Other hydrogen bonds between cathepsin B and E-64 are summarized in Table 1. The cathepsin B-leupeptin complex demonstrated no hydrogen bonding to the active site amino acids (Figure 10-a), while there was a hydrogen bond between the nitrogen atom 4 of glycine 60 cathepsin B with the oxygen atom 10 of leupeptin. It also revealed a hydrogen bond between the oxygen atom 8 of the asparagine 151 of cathepsin B and the nitrogen atom 3 of leupeptin ligand (Figure 10-b).

According to Figure 11a, the cathepsin B-cystatin C complex (25KQIVAGVNYfLDVE35) suggested no hydrogen bonding to the cathepsin B active site amino

**Table 1.** Hydrogen Bond (HB) and Electrostatic Interaction (EI)

| Complex Interaction | Hydrogen Bond | Electrostatic Interaction | Distance (Å) |
|---|---|---|---|
| Cathepsin B_E-64<br>Complex Energy (-71 kcal/mol) | Gly 64 H_E-64 O | | 1.2 |
| | Lys 67 H_E-64 O | | 1.2 |
| | Cys 108 O_E-64 N | | 1.52 |
| | Gly 112 N_E-64 N | | 1.55 |
| | Gly 251 O_E-64 N | | 1.52 |
| | Ser 299 O_E-64 N | | 1.52 |
| | Glu 324 O_E-64 O | | 1.52 |
| | | Glu 115 H_E-64 N (2) | 1.2 |
| | | His 278 O_E-64 N | 1.2 |
| Cathepsin B_Leupeptin<br>Complex Energy (-28 kcal/mol) | Ans 16 N_Leu O | | 1.55 |
| | Gly 60 N_Leu O | | 1.55 |
| | Thr 61 O_Leu N | | 1.52 |
| | Gly 277 N_Leu O | | 1.55 |
| Cathepsin B_Cystatine C<br>Complex Energy (-37 kcal/mol) | Arg 18 H_Cys Glu 26 O (2) | | 1.52 |
| | | Asp 184 O_Cys Gly 30 N | 1.52 |



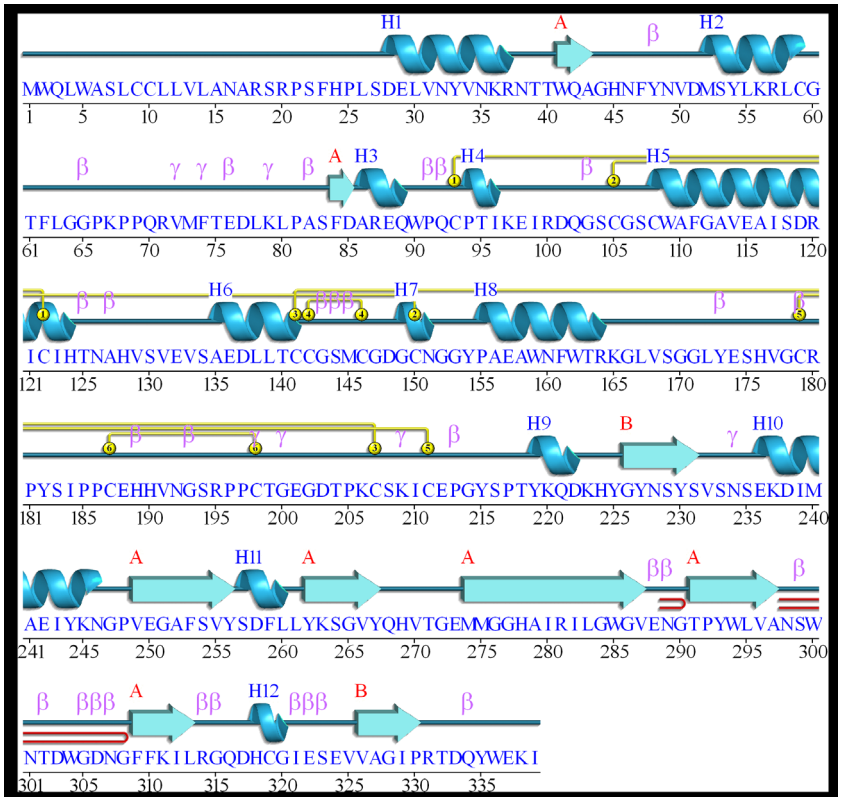
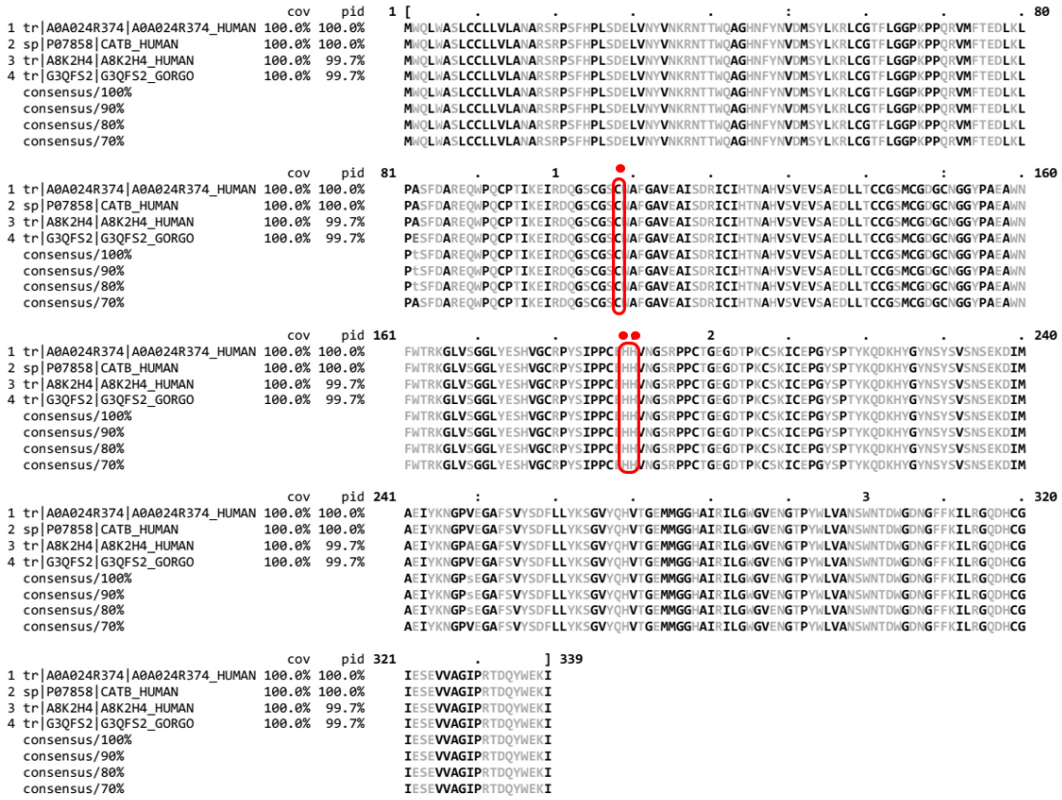





**Figure 1.** The second structure of cathepsin B

(  ): Alpha-helix, beta-sheet, and beta-turn, respectively.

(  ): The probability of disulfide bond formation.




International Journal of
Medical Toxicology & Forensic Medicine

**Figure 2.** The cathepsin B active site amino acids, compared to other similar proteins

The cysteine 108, histidine 189, and histidine 190 amino acids are located in conserved regions. The red circle (•) indicates the amino acids of the active site.

acids. None of the amino acids active site of cathepsin B hydrogen-bonded with cystatin C; however, the oxygen atom 8 of the arginine 18 of cathepsin B, formed a hydrogen bond with the oxygen atom 16 of the glutamine 26 of cystatin. Furthermore, the oxygen atom 8 of aspartate 148 of cathepsin B formed an electrostatic bonded with the nitrogen atom 6 of glycine 30 of the cystatin C (Figure 11b).

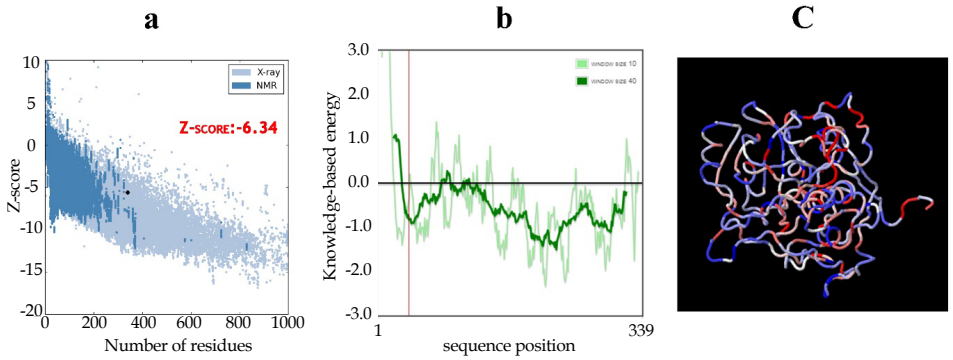



**Figure 3.** Predicted modeling homology and the validation of cathepsin B structure

International Journal of
Medical Toxicology & Forensic Medicine

a. The Z-score graph of ProSA; b. Local model quality in this graph, the amino acids are compared to 40 amino acids and 10 amino acids, respectively; if the value is below zero, it indicates the high quality of the predicted model; c. Three-dimensional structures of proteins energy; the blue areas indicate the low energy level of the predicted model.

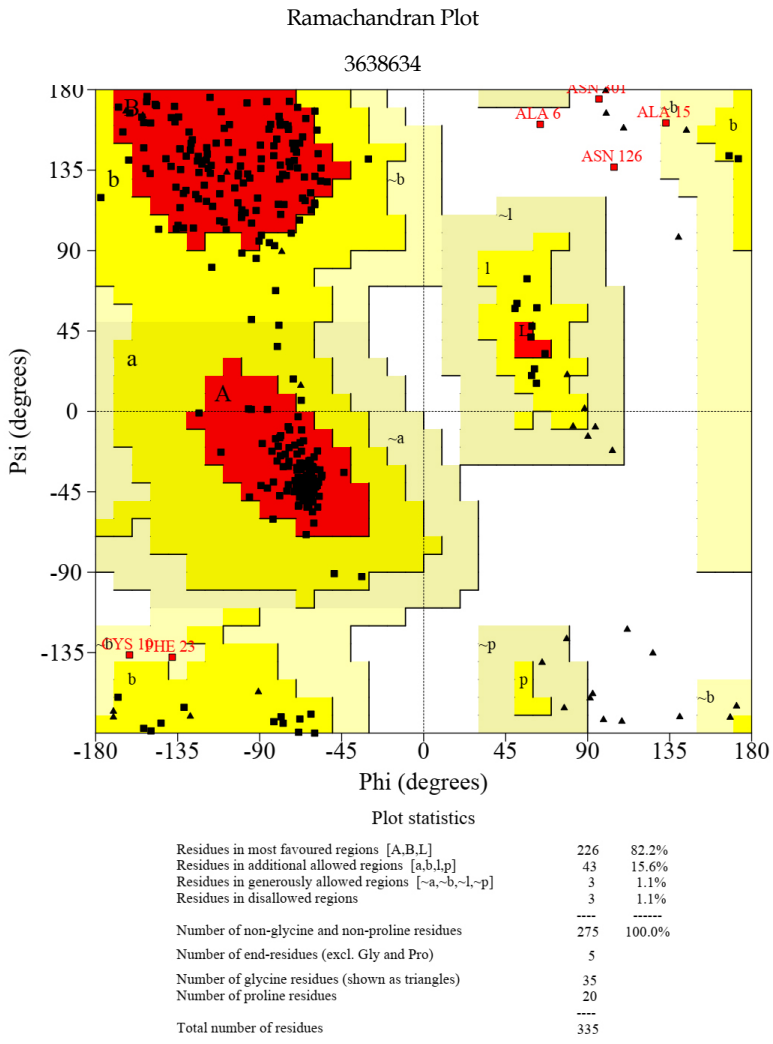



**Figure 4.** Ramachandran graph

International Journal of Medical Toxicology & Forensic Medicine

In total, 99% of the amino acids in the predicted three-dimensional structure are correctly angled in terms of the Phi (φ) and Psi (ψ) angles.

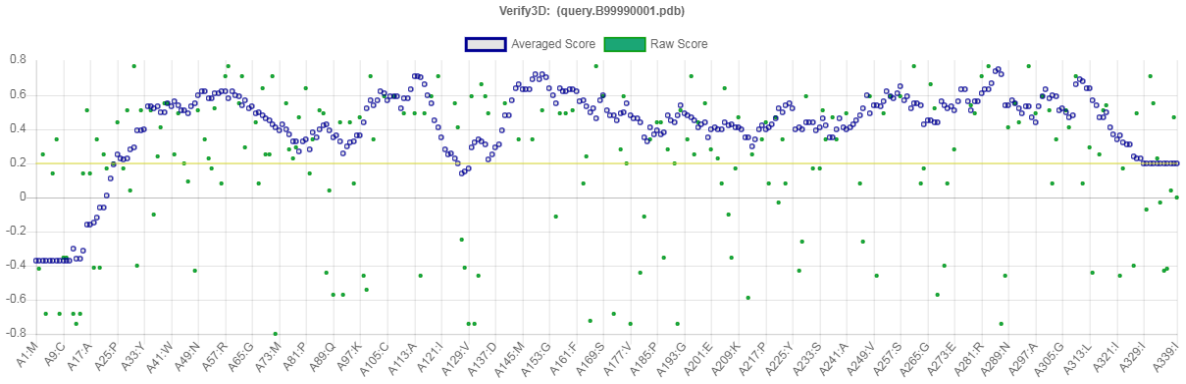



**Figure 5.** VERIFY-3D graph

International Journal of Medical Toxicology & Forensic Medicine

Compatibility of cathepsin B atomic model of the 3D structure with the cathepsin B primary structure

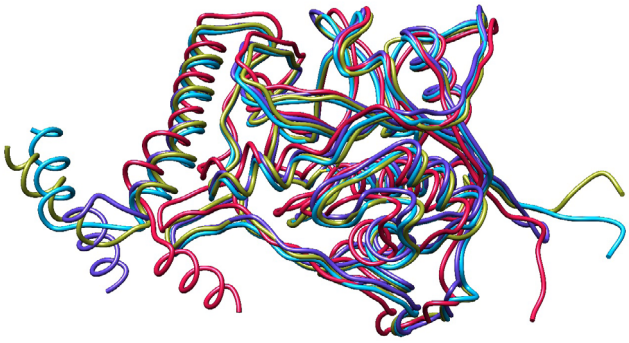




**Figure 6.** Comparing the predicted 3D structure of cathepsin B by MODELLER with the predicted 3D structure of I-Tasser

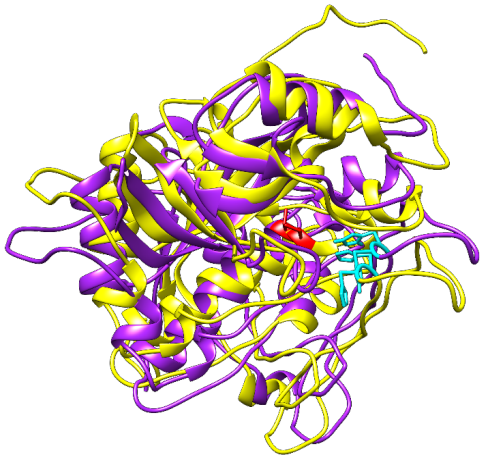




**Figure 7.** Adaptation the predicted 3D structure of cathepsin B (Amino acid 1 to 339) and the crystallographic 3D structure of 1PBH (Amino acid 80 to 420) using Superimpose and comparing the positioning of active site amino acids in both models

Yellow color: Cathepsin B predicted 3D structure; Purple color: 1PBH 3D Structure; Red color: Cysteine amino acid, Turquoise color: Histidine amino acid

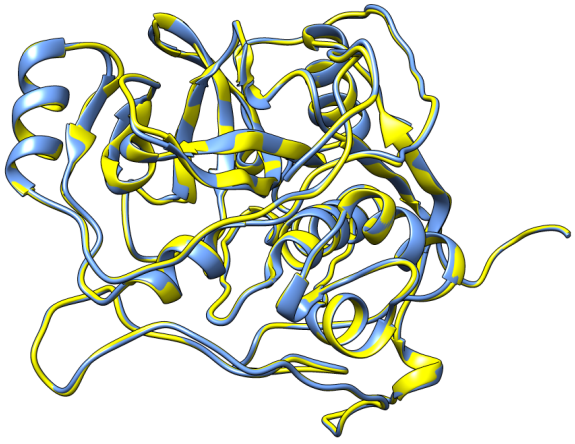




**Figure 8.** Superimpose shape of cathepsin B predicted 3D structure in the minimized state with its 3D structure

Yellow color: Cathepsin B predicted 3D structure; Turquoise color: Cathepsin B predicted 3D structure in minimized mode

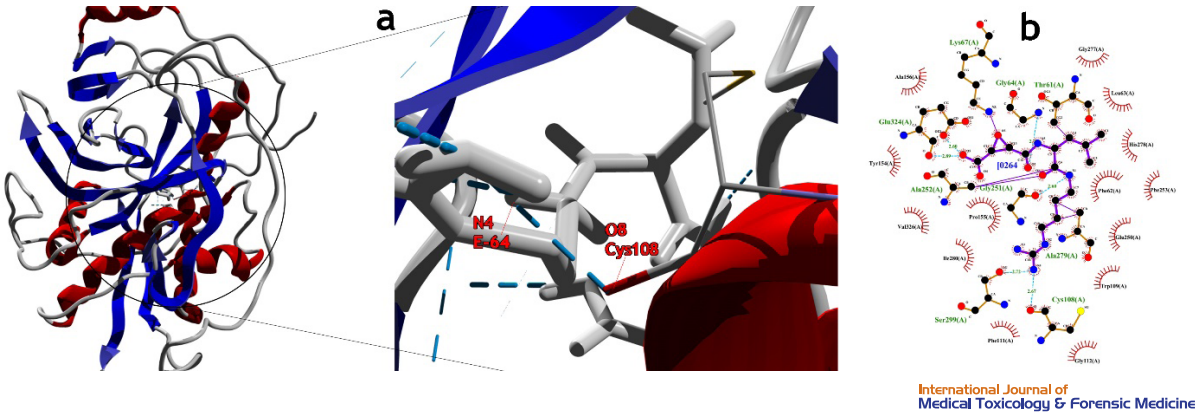



**Figure 9.** Cathepsin B _E-64 complex docking results

a. Cathepsin B _E-64 complex; Interaction between the cysteine amino acid 108 active sites of cathepsin B and E-64 Thin red color: Cysteine amino acid; wide white color: E-64; cross-hatching blue: Hydrogen bond; b. 2D graph of the amino acids involved in the cathepsin B_E-64 complex; Green color: Hydrogen bonded amino acids; Black color: hydrophobic amino acids with electrostatic and stearic bonds

## 4. Discussion

Cathepsins are considered as a class of tumor markers. The extent and location of expression of B and L cathepsins in tumors vary with normal tissues [7]. Increased expression of cathepsin B has been observed in various cancers, including breast, colon, stomach, lung, and prostate types [23]. The imbalance between cathepsins and their endogenous inhibitors has been implicated in the processes of cancer formation, tumor invasion, and metastasis [24]. Various studies have examined the role of cystatin C as one of the major inhibitors of cathepsins. For example, the growth of lymph cancer cells in mice was associated with decreased extracellular cystatin C content in the tumor and plasma.

Moreover, antitumor drug treatment led to increased lifespan, reduced tumor size, and increased cystatin C content in tumor and plasma tissues [25]. Studies have reported that measuring cystatin C in tumor and plasma tissues could help predict tumor growth and evaluate the efficacy of antitumor treatment [26]. In a study, the effect of cystatin C on papain, cathepsin B, H, and L was investigated and found to decrease or completely inhibit their activity [27]. Numerous in vitro studies have been performed to investigate mechanisms that control and reduce the expression and inhibition of cathepsin B activity [28]. In addition to natural inhibitors, like cystatin C, chemical inhibitors, such as E-64 and leupeptin, could also inhibit cathepsin B activity [9, 29]. A study assessed the inhibition of cathepsin B activity by E-64 induction [30]. In vitro, leupeptin also reduced the activity of cathepsin B [31, 32]. Yoko Hashimoto identified the vari-

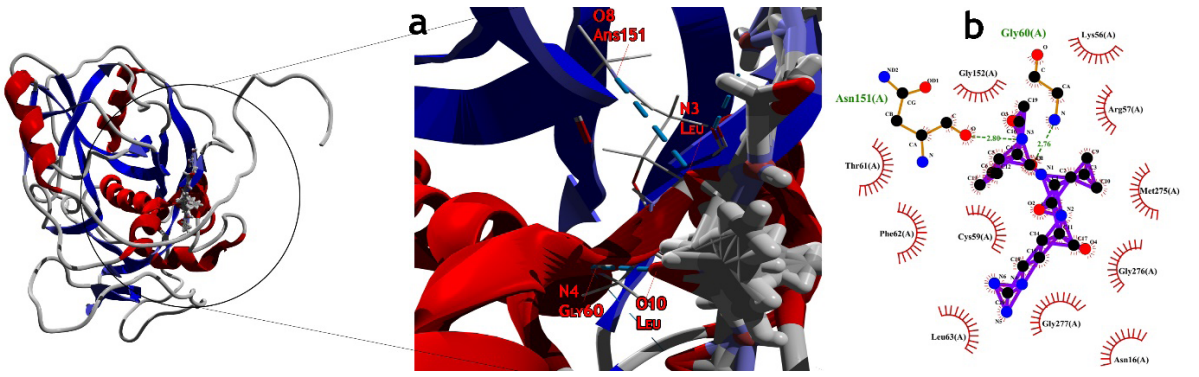



**Figure 10.** Cathepsin B_Leupeptin Complex docking results

a. Cathepsin B_Leupeptin Complex; Interaction between glycine 60 and asparagine 151 amino acids of cathepsin B with Leupeptin; Thin red color: glycine 60 and asparagine 151 amino acid; wide white color: Leupeptin; cross-hatching blue: Hydrogen bond; b. 2D graph of the amino acids involved in the cathepsin B_Leupeptin complex; Green color: Hydrogen bonded amino acids; Black color: hydrophobic amino acids with electrostatic and stearic bonds

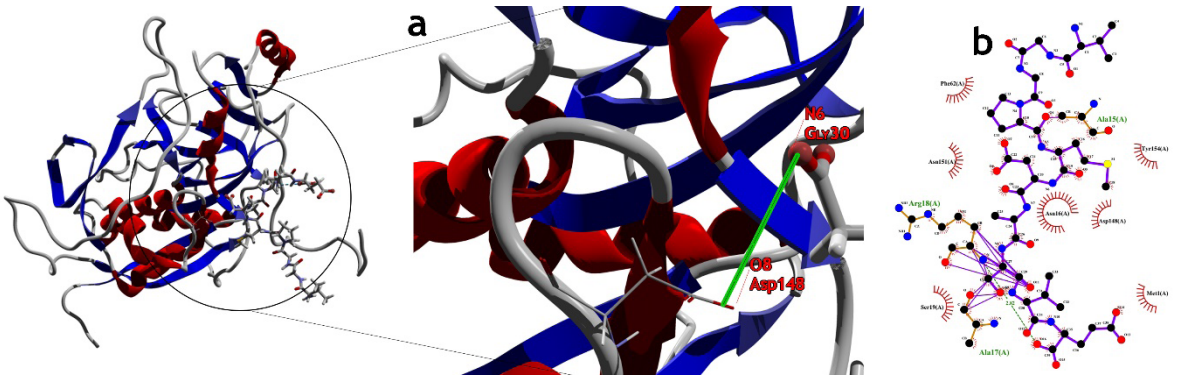



**Figure 11.** Cathepsin B_Leupeptin Complex docking results



a. Cathepsin B-Cystatin C Complex; Interaction between aspartate 148 of cathepsin B with glycine 30 cystatin C. Thin red color: aspartate amino acid, wide white color: glycine amino acid, continuous green: electrostatic bond. b) 2D graph of the amino acids involved in the cathepsin B_ cystatin C complex. Green color: Hydrogen bonded amino acids, Black color: hydrophobic amino acids with electrostatic and stearic bonds.

ants of cathepsin L by zymogram using the inhibitory property of leupeptin [33].

The evaluation and prediction of the activation mechanism of cathepsin B inhibitors using bioinformatics software are among the most accurate and cost-effective methods in drug design and cancer suppressors. The in vitro use of inhibitors requires time, cost, and equipment. In this study, the interaction between cathepsin B and inhibitors was detected using bioinformatics software. First, the 3D modeling of cathepsin B was performed. Then, the interaction between the predicted model of cathepsin B with cystatin C, E-64 and leupeptin were investigated.

Based on bioinformatics databases, cathepsin is a class of cysteine protease family. Cathepsin B has a domain with one cysteine amino acid 108 and two histidine amino acids 189 and 190 in the active site; i.e., nucleophilic attacked by the cysteine amino acid present in the active site. Bioinformatics studies determined some biochemical properties of cathepsin B, the presence of signal peptide, and the formation of disulfide bonds. The achieved instability index was 31.25, i.e., in the range of stable enzymes with a numerical range of <40. The aliphatic coefficient of the enzyme, which indicates the volume occupied by the side chain of amino acids, ALa, Val, Leu, and Ile, was equal to 64.69.

ProFunc and PSIPRED databases were used to predict the secondary structure of the protein. The Self-Optimized Prediction Method With Alignment (SOPMA) analysis results revealed that alpha-helix constituted the largest amount of cathepsin B secondary structure. Comparing I-Tasser's active site prediction with other cathepsin B was used for molecular modeling and docking studies.

Since the cathepsin B crystallographic sequence exists only as procathepsin, modeling was performed from sequence 1 to 339 of cathepsin B amino acids. Predictive model validation was performed by online software PROCHECK, PROSA, and GENO3D. The Z-SCORE point was set at -6.34. The Z-score is used to evaluate the validity of the modeled protein based on Nuclear Magnetic Resonance (NMR) or X-ray methods. The NMR method is used for the crystallography of proteins with <200 amino acids, i.e., more accurate than X-ray.

If the Z-Score dot lies in the NMR or X-Ray region, the simulation accuracy is higher; thus, the simulated model has the lowest error rate and the highest confidence. Furthermore, in the graph plotted to determine the negative energy level of the protein if the amino acid sequence is found to be 10 and 40 less than zero, the low energy level signifies greater protein stability and the high value of the simulated 3D data for the protein.

In the Ramachandran graph, amino acids are grouped according to the Phi (φ) and Psi (ψ) angles. The final quality of the predicted model is acceptable, because generally, 99% of the amino acids are in their proper place. The VERIFY3D program determines the compatibility of the atomic model of cathepsin B predicted 3D with its first amino acid sequence. The compatibility scores above zero in the VERIFY3D graph are related to the adaptation degree of the cathepsin B amino acid sequence in this pattern. The VERIFY3D results for cathepsin B were >2; therefore, the predicted model was associated with high affinities, compared to the 1BPH

structure. These results indicate that the homology modeling of the predicted cathepsin B structure has been successfully accomplished; accordingly, this model is beneficial for applying in docking reviews.

Superimposition using Chimera (RMSD Å 2.32) revealed the active site of the cathepsin B structure. Such structure consisted of the amino acids Cys 108, His 189, His 190, i.e, very similar to the active site of the procathepsin 1PBH structure (80 to 420 amino acid sequences); the amino acids were Cys 32, His 112, His 113. With Chimera software, the best spatial positioning for optimal enzyme energy in the dock was achieved. This software calculates and estimates the total energy of the predicted structure of cathepsin B, making the protein ideally stable and reducing energy.

Cathepsin B belongs to the cysteine proteases group, the main characteristic of cysteine proteases being the presence of cysteine and histidine amino acids in the enzyme's active site. The cysteine amino acid of the active site in cathepsin B significantly affects the nucleophilic attack. Thus, it is important to investigate compounds that could inhibit or interfere with this site's function. Therefore, we investigated the interaction of cathepsin B with cystatin C, as a natural inhibitor, as well as the interaction of cathepsin B with E-64 and leupeptin, as synthetic compounds.

Molecular docking studies were used to understand the protein-ligand and protein-protein interaction in this study. The docking of the protein-ligand was performed between the cathepsin's active site, and E-64 and leupeptin. Protein-protein docking was also established between cathepsin B and cystatin C. The cystatin PDB structure was downloaded from the RCSB database, and initial preparation for docking was performed by MVD. The structures of E-64 and leupeptin as ligands obtained from PubChem were also evaluated for energy and stability and prepared for docking.

Applying E-64 in the in vivo environment is due to its high permeability to cells and tissues, as well as its low toxicity. As noted in previous research, E-64, through its trans-epoxy- succinyl group, forms an irreversible bond with the cysteine amino thiol group of the active site of cysteine protease. This process ultimately inhibits cathepsin B activity by forming a linkage. Atsushi Yamamoto implemented a specific method and argued that E-64 interacts with the cathepsin B active site of the cysteine 32amino acid [34]. This result, consistent with that of the present study, demonstrates the interaction of E-64 with the active site of cathepsin B. Although leupeptin, like E-64, is an epoxide compound, the cathepsin B-leupeptin complex demonstrated no hydrogen bond in the active site amino acids.

As per Figure 11a, the cathepsin B-cystatin C complex (25KQIVAGVNYfL DVE35) revealed no hydrogen bonding to the active site amino acids. A study explored the interaction of cathepsin L with a cystatin family. As a result, it was reported that no hydrogen bonds were established between cathepsin L and cystatin C; however, an electrostatic bond was established between cathepsin L and cystatin C. We analyzed the results of cathepsin L docking with the A and B Stefin groups, as well as the cystatin groups C, D, F, M, E, S, SA and SN.

Accordingly, it was found that the two amino acids of ANS66 and A SP162 derived from cathepsin L, interacted with all inhibitors in combination [35]. Since the active site was similar to cathepsin B and L, these results were consistent with the molecular docking result of our study. In other words, in our study, an electrostatic bond was established between the aspartate 148 of cathepsin B and glycine 30 of cystatin C. These results indicated that our research using bioinformatics software was largely consistent with the experimental results.

## 5. Conclusion

Cathepsin B, as a cysteine protease, plays a highly beneficial role in the intracellular processes of protein catabolism; however, under the conditions of cancer, it plays a very positive role in furthering cancer therapy. In cancer, cathepsin B negatively affects the metabolic process of cells by degrading and destroying the extracellular matrix and stimulates the cancer metastasis stage. Additionally, by interfering with the body's immune resistance, it promotes the proper growth of malignant tumors. As the cancer progresses, cathepsin B reduces cystatin C expression.

Decreased cystatin C expression level disrupts the balance of cathepsin B/cystatin C ratio; finally, a decrease in cystatin C concentration, as a natural inhibitor of cysteine proteases, reduces its ability to inhibit cathepsin B activity. In this study, we could construct the cathepsin B 3D model and ensure the predicted model's quality. After docking cathepsin B with three cysteine protease inhibitors, it was found that cystatin C forms a strong electrostatic bond with the aspartate 148 of cathepsin B. However, the E-64, and leupeptin, as two epoxide inhibitors for cysteine proteases, E-64 could form a hydrogen bond with the active site cysteine 108 carboxylic group of cathepsin B. This interaction may reduce the ability

of the cysteine 108 amino acid sulfur atom to attack the nucleophilic peptide bond. Therefore, investigating the interaction by docking in this study could help to select the type of inhibitor and identify the methods of cancer control.

## Ethical Considerations

### Compliance with ethical guidelines

All ethical principles were considered in this article.

### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Author's contributions

Conceptualization, Writing Review & Editing, Funding Acquisition, and Supervision: Afshin Khara, Ehsan Jahangirian; Methodology and Writing: Ehsan Jahangirian, Hossein Tarrahimofrad; Investigation: Afshin Khara, Ehsan Jahangirian, Hossein Tarrahimofrad; Resources: Ehsan Jahangirian;Supervision: Afshin Khara, Ehsan Jahangirian.

### Conflict of interest

The authors declared no conflict of interest.

## References

[1] Turk V, Stoka V, Vasiljeva O, Renko M, Sun T, Turk B, et al. Cysteine cathepsins: From structure, function and regulation to new frontiers. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 2012; 1824(1):68-88. [DOI:10.1016/j.bbapap.2011.10.002] [PMID]

[2] Calkins CC, Sameni M, Koblinski J, Sloane BF, Moin K. Differential localization of cysteine protease inhibitors and a target cysteine protease, cathepsin B, by immuno-confocal microscopy. The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society. 1998; 46(6):745-51. [DOI:10.1177/002215549804600607] [PMID]

[3] Bell-McGuinn KM, Garfall AL, Bogyo M, Hanahan D, Joyce JA. Inhibition of cysteine cathepsin protease activity enhances chemotherapy regimens by decreasing tumor growth and invasiveness in a mouse model of multistage cancer. Cancer Research. 2007; 67(15):7378-85. [DOI:10.1158/0008-5472.CAN-07-0602] [PMID]

[4] Fonović M, Turk B. Cysteine cathepsins and their potential in clinical therapy and biomarker discovery. Proteomics. Clinical Applications. 2014; 8(5-6):416-26. [DOI:10.1002/prca.201300085] [PMID]

[5] Turk V, Stoka V, Turk D. Cystatins: Biochemical and structural properties, and medical relevance. Frontiers in Bioscience: A Journal and Virtual Library. 2008; 13:5406-20. [DOI:10.2741/3089] [PMID]

[6] Sokol JP, Schiemann WP. Cystatin C antagonizes transforming growth factor beta signaling in normal and cancer cells. Molecular Cancer Research. 2004; 2(3):183-95. [PMID]

[7] Goulet B, Sansregret L, Leduy L, Bogyo M, Weber E, Chauhan SS, et al. Increased expression and activity of nuclear cathepsin L in cancer cells suggests a novel mechanism of cell transformation. Molecular Cancer Research. 2007; 5(9):899-907. [DOI:10.1158/1541-7786.MCR-07-0160] [PMID]

[8] Barrett AJ, Kembhavi AA, Brown MA, Kirschke H, Knight CG, Tamai M, et al. L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. The Biochemical Journal. 1982; 201(1):189-98. [DOI:10.1042/bj2010189] [PMID] [PMCID]

[9] Katunuma N, Kominami E. Structure, properties, mechanisms, and assays of cysteine protease inhibitors: Cystatins and E-64 derivatives. Methods in Enzymology. 1995; 251:382-97. [DOI:10.1016/0076-6879(95)51142-3]

[10] Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the expasy server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. [DOI:10.1385/1-59259-890-0:571]

[11] Dyrløv Bendtsen J, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. Journal of Molecular Biology. 2004; 340(4):783-95. [DOI:10.1016/j.jmb.2004.05.028] [PMID]

[12] Ceroni A, Passerini A, Vullo A, Frasconi P. DISULFIND: A disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Research. 2006; 34(Web Server Issue):W177-81. [DOI:10.1093/nar/gkl266] [PMID] [PMCID]

[13] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 2015; 10:845. [DOI:10.1038/nprot.2015.053] [PMID] [PMCID]

[14] Wass MN, Kelley LA, Sternberg MJ. 3DLigandSite: Predicting ligand-binding sites using similar structures. Nucleic Acids Research. 2010; 38(Web Server Issue):W469-73. [DOI:10.1093/nar/gkq406] [PMID] [PMCID]

[15] Zhang C, Freddolino PL, Zhang Y. COFACTOR: Improved protein function prediction by combining structure, sequence and protein-protein interaction information. Nucleic Acids Research. 2017; 45(W1):W291-9. [DOI:10.1093/nar/gkx366] [PMID] [PMCID]

[16] Roy A, Yang J, Zhang Y. COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research. 2012; 40(Web Server Issue):W471-7. [DOI:10.1093/nar/gks372] [PMID] [PMCID]

[17] Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. Journal of Molecular Biology. 1996; 264(1):121-36. [DOI:10.1006/jmbi.1996.0628] [PMID]

[18] Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional struc-

tures of proteins. Nucleic Acids Research. 2007; 35(Suppl. 2):W407-10. [DOI:10.1093/nar/gkm290] [PMID] [PMCID]

[19] Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins: Structure, Function, and Bioinformatics. 1993; 17(4):355-62. [DOI:10.1002/prot.340170404] [PMID]

[20] Depristo MA, de Bakker PI, Johnson RJ, Blundell TL. Crystallographic refinement by knowledge-based exploration of complex energy landscapes. Structure. 2005; 13(9):1311-9. [DOI:10.1016/j.str.2005.06.008] [PMID]

[21] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF chimera--a visualization system for exploratory research and analysis. Journal of Computational Chemistry. 2004; 25(13):1605-12. [DOI:10.1002/jcc.20084] [PMID]

[22] Laskowski RA, Swindells MB. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. Journal of Chemical Information and Modeling. 2011; 51(10):2778-86. [DOI:10.1021/ci200227u] [PMID]

[23] Mohamed MM, Sloane BF. Multifunctional enzymes in cancer. Nature Reviews Cancer. 2006; 6(10):764-75. [DOI:10.1038/nrc1949] [PMID]

[24] Szpaderska AM, Frankfater A. An intracellular form of cathepsin B contributes to invasiveness in cancer. Cancer Research. 2001; 61(8):3493-500. [PMID]

[25] Kopitz C, Anton M, Gansbacher B, Krüger A. Reduction of experimental human fibrosarcoma lung metastasis in mice by adenovirus-mediated cystatin C overexpression in the host. Cancer Research. 2005; 65(19):8608-12. [DOI:10.1158/0008-5472.CAN-05-1572] [PMID]

[26] Palermo C, Joyce JA. Cysteine cathepsin proteases as pharmacological targets in cancer. Trends in Pharmacological Sciences. 2008; 29(1):22-8. [DOI:10.1016/j.tips.2007.10.011] [PMID]

[27] Cimerman N, Prebanda MT, Turk B, Popovič T, Dolenc I, Turk V. Interaction of cystatin C variants with papain and human cathepsins B, H and L. Journal of Enzyme Inhibition. 1999; 14(2):167-74. [DOI:10.3109/14756369909036552] [PMID]

[28] Costantino CM, Ploegh HL, Hafler DA. Cathepsin S regulates class II MHC processing in human CD4+ HLA-DR+ T cells. The Journal of Immunology. 2009; 183(2):945-52. [DOI:10.4049/jimmunol.0900921] [PMID] [PMCID]

[29] Saino T, Someno T, Ishii S, Aoyagi T, Umezawa H. Protease-inhibitory activities of leupeptin analogues. The Journal of Antibiotics. 1988; 41(2):220-5. [DOI:10.7164/antibiotics.41.220] [PMID]

[30] Wadhawan M, Singh N, Rathaur S. Inhibition of cathepsin B by E-64 induces oxidative stress and apoptosis in filarial parasite. PLOS ONE. 2014; 9(3):e93161. [DOI:10.1371/journal.pone.0093161] [PMID] [PMCID]

[31] Sutherland JH, Greenbaum LM. Paradoxical effect of leupeptin in vivo on cathepsin B activity. Biochemical and Biophysical Research Communications. 1983; 110(1):332-8. [DOI:10.1016/0006-291X(83)91300-1]

[32] Baici A, Gyger-Marazzi M. The slow, tight-binding inhibition of cathepsin B by leupeptin. A hysteretic effect. European Journal of Biochemistry. 1982; 129(1):33-41. [DOI:10.1111/j.1432-1033.1982.tb07017.x] [PMID]

[33] Hashimoto Y. Gelatin zymography using leupeptin for the detection of various cathepsin l forms. Methods in Molecular Biology (Clifton, NJ). 2017; 1594:243-54. [DOI:10.1007/978-1-4939-6934-0_16] [PMID]

[34] Yamamoto A, Tomoo K, Matsugi K, Hara T, In Y, Murata M, et al. Structural basis for development of cathepsin B-specific noncovalent-type inhibitor: Crystal structure of cathepsin B-E64c complex. Biochimica et Biophysica Acta. 2002; 1597(2):244-51. [DOI:10.1016/S0167-4838(02)00284-4]

[35] Nandy SK, Seal A. Structural dynamics investigation of human family 1 & 2 cystatin-cathepsin l1 interaction: A comparison of binding modes. PLOS ONE. 2016; 11(10):e0164970. [DOI:10.1371/journal.pone.0164970] [PMID] [PMCID]